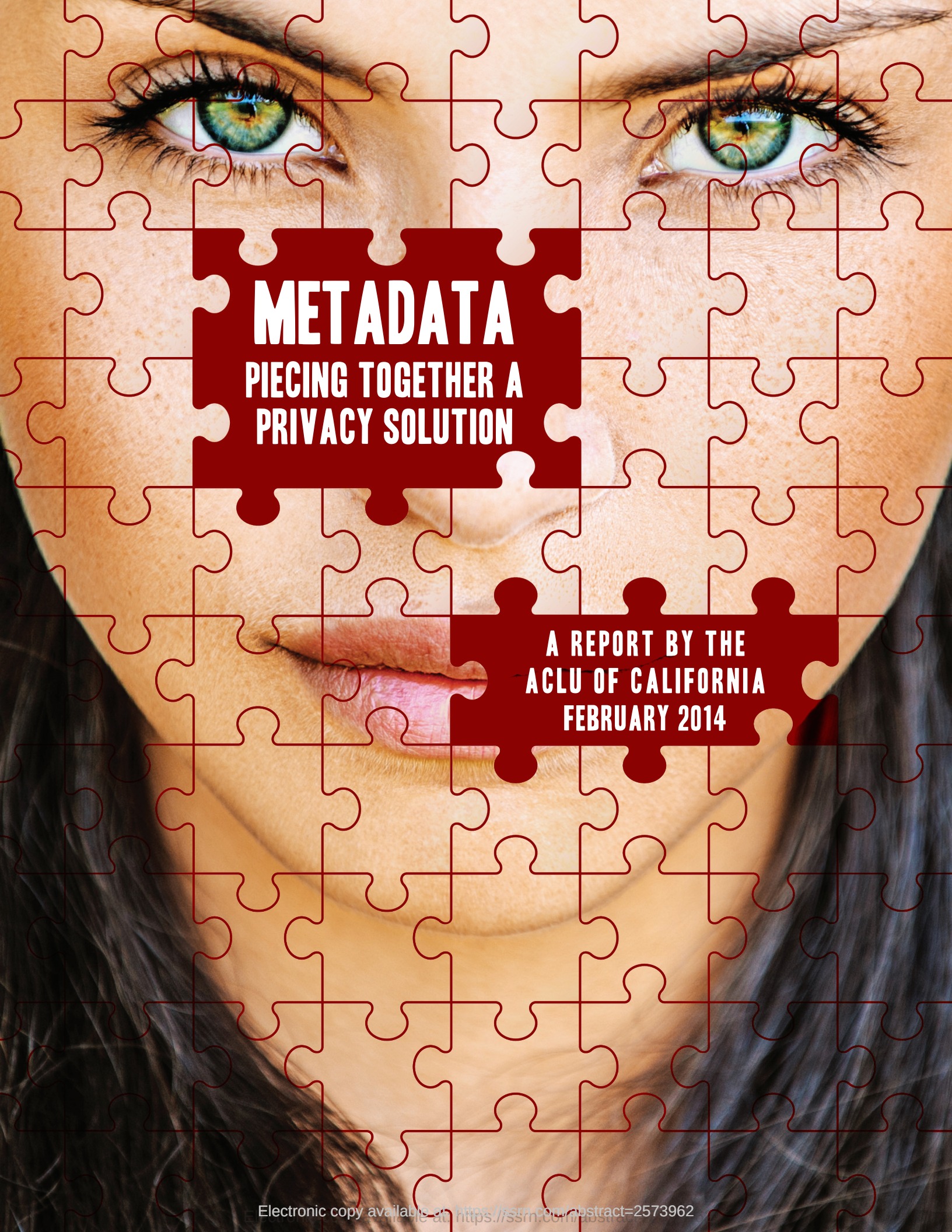
2-2014

Metadata: Piecing Together a Privacy Solution

Chris Conley



Imagine the government is constantly monitoring you — keeping track of every person you call or email, every place you go, everything you buy, and more — all without getting a warrant. And when you challenge them, they claim you have no right to expect this kind of information to be private. Besides, they're not actually listening to what you say or reading what you write, so what's the big deal anyhow?

Unfortunately, this scenario is more real than imaginary. Government agencies ranging from the NSA to local police departments have taken advantage of weak or uncertain legal protections for **"metadata"** — descriptive information about our phone calls, emails, location, purchases, and more — to sweep up vast amounts of information about innocent Americans without a warrant.

Limited privacy protections for metadata may have made sense decades ago when technology to collect and analyze data was virtually nonexistent. But in today's "big data" world, non-content does not mean non-sensitive. In fact, new technology is demonstrating just how sensitive metadata can be: how friend lists can reveal a person's sexual orientation, purchase histories can identify a pregnancy before any visible signs appear, and location information can expose individuals to harassment for unpopular political views or even theft and physical harm.

Two separate committees assembled by the executive branch — the President's Review Group on Intelligence and Communications Technology and the Privacy and Civil Liberties Oversight Board —have joined lawmakers, academics, and judges in calling for a reevaluation of the distinction between content and metadata. This paper examines how new technologies and outdated laws have combined to make metadata more important and more vulnerable than ever, and proposes a way forward to ensure that all of our sensitive information gets the privacy protection it deserves.

For more information, please visit **www.aclunc.org/tech/meta**.

Electronic copy available at: https://ssrn.com/abstract=2573962

TABLE OF CONTENTS

AUTHOR: Chris Conley, Technology and Civil Liberties Project, ACLU of Northern California

DESIGN: Anna Salem, ACLU of Northern California

COVER: Gigi Pandian, ACLU of Northern California

PRINTING: Inkworks Press

PUBLISHED BY: ACLU of California, February 2014

THANKS TO: Nicole Ozer and Matt Cagle, ACLU of Northern California Technology and Civil Liberties Project, and attorneys and staff of the National ACLU's Center for Democracy, the Speech, Privacy, and Technology Project, and the Washington Legislative Office for contributions to and feedback on this report.

This publication is supported by cy pres funds and the generosity of the ACLU's members and donors.

Electronic copy available at: https://ssrn.com/abstract=2573962

INTRODUCTION

Imagine bringing a date home for dinner. You put the laptop away and mute your phone. You prepare a gourmet home-cooked meal for two, queue up a selection of romantic songs, and pick out a movie to watch after dinner. As the evening winds down, your heart races a bit as you go in for a kiss and wonder how your night will end.

Now imagine that someone is monitoring each and every event of your evening. Oh, don't worry, they're not actually watching you or listening in on your conversation. They just know who you emailed or called just before you put your computer away. They know what you bought for dinner and how you prepared it. They know who came over, where he or she came from, and how long he or she stayed. They know what time you started the movie and which songs you listened to. They even know what time you turned off the lights — and whether or not the music was still playing when you did. And they know all of this without ever getting a search warrant.

Can they do that?

Unfortunately, the President, the National Security Agency (NSA), and federal and state law enforcement agencies seem to believe the answer is yes.

Generally speaking, the government cannot record or obtain the contents of your communications without at least a search warrant. But "metadata," information other than communications content, is often treated differently under the law. As a result, government entities ranging from local police departments to the NSA have asserted broad authority to acquire location information, associational data, records of purchases and financial transactions, and more, all of which can reveal intimate details of your life.

In the modern world, non-content does not mean non-sensitive.

And the government doesn't merely believe that it can collect this kind of information without a search warrant — it is actually doing so, in quantities that beggar the imagination. Recent revelations about NSA surveillance programs reveal that the agency has attempted to obtain information about every phone call and every Internet communication carried by U.S. networks. Law enforcement agencies have tracked the location of individuals for months at a time. Telecommunications companies have revealed that they respond to millions of requests for information from law enforcement agencies at all levels, and at least one carrier provides the Drug Enforcement Agency (DEA) with access to records about every call that has crossed the carrier's network since 1987. Yet despite the obvious privacy concerns that arise from these practices, all of this regularly occurs without the judicial oversight and checks and balances that the Constitution requires.

Electronic copy available at: https://ssrn.com/abstract=2573962

In response to disclosures about the NSA, President Obama told the American people not to worry: "We're not listening in on your calls." But that is poor solace once we realize what the government is doing: recording our presence at a hospital, a political rally, or a religious ceremony; tracking our calls to an addiction support hotline, a job recruiter, or a dating service; monitoring our purchases of birth control or books on fighting depression; and far more. These details about our physical and mental health, business and personal associations, and financial and relationship status have been swept into databases where they have been improperly accessed thousands of times. All this because decades-old statutes and court decisions supposedly draw a line that wraps protection around content but leaves everything else out in the cold.

"We recommend that the government should commission a study of the legal and policy options for assessing the distinction between meta-data and other types of information."

President's Review Group on Intelligence and Communications Technologies

This distinction between content and metadata might have made sense decades ago when technology to collect and analyze data was virtually nonexistent. But in the modern world, non-content does not mean non-sensitive. In fact, the explosion of data mining, targeted advertising, and other new technologies is driven by the realization that companies and the government can learn a great deal about an individual simply by recording his or her actions. Lawmakers, judges, academics, and two separate committees assembled by the executive branch (the President's Review Group on Intelligence and Communications Technology and the Privacy and Civil Liberties Oversight Board) have all recognized the need to reevaluate the distinction between content and metadata.

This paper contends that, in light of extensive evidence of the potential and actual abuse of metadata collection, metadata should generally receive the same legal protection as the contents of a communication. Part I of this paper attempts to define exactly what metadata means, highlighting the semantic and contextual difficulties with the term itself, and examines the wide array of information that can be exposed through metadata. Part II highlights the abuses and questionable practices that have occurred where metadata has received inadequate privacy protections. Part III looks at recent efforts by policymakers to enhance privacy protections for at least some types of metadata. Finally, Part IV concludes that the general distinction between metadata and content is no longer valid in the modern world and lays out a roadmap to establish strong privacy protections for all sensitive information.

As Part I illustrates, in some cases it is not even easy to determine whether some particular piece of information is metadata at all. But the effort to do so misses the point. What truly matters is that the law protects an individual's sensitive personal information, whether that information is contained within the contents of a communication or is exposed through the collection or analysis of metadata.

Electronic copy available at: https://ssrn.com/abstract=2573962

PART I: What Is Metadata, and Why Is It Important?

For purposes of this paper, we use the term "metadata" to mean "any data other than the contents of a communication" — which, of course, leads to the natural follow-up question: what is content? Unfortunately, answering that "persnickety, but persistent question"[1] is far more challenging than it might first appear. In many cases, there is no distinct boundary between content and metadata; in others, precisely the same information may be content in one context and metadata in another.

What is clear, however, is that even drawing the circle around content as widely as possible leaves an incredible amount of sensitive personal information outside of its boundaries. Information about our location derived from cell phone towers, about our associations drawn from the To and From lines of our emails, about the items we purchase, and more may not be the contents of our communications, but they can paint a profoundly detailed picture of our lives. As a result, misusing or inappropriately disclosing information derived from metadata can significantly harm both individuals and our democratic principles. As a result President Obama's reassurance that "we're not listening to your calls" rings hollow — why should that console us if the exact same information can in so many cases be extracted from our metadata?

Defining Metadata

Metadata was originally understood to mean "data about data."[2] Its modern connotation is considerably broader, encompassing a wide range of information ranging from traditional metadata (e.g., the time a file was created) to records of interactions (a history of login times) to simple facts about individuals (an account number or mailing address). For the most part, metadata is defined by what it is not: the contents of a communication. Unfortunately, that simply shifts the question from defining metadata to defining content.

One technical approach to distinguishing content and metadata is to divide electronic communications into "payload" (content) and "delivery instructions" (metadata). The body of an email you write, the photo you upload, and the comment you post on a social network are the substance of what you are communicating, and therefore content. On the other hand, the address to which you sent that email, the account to which you uploaded the photo, or the IP address of the computer you use to post that comment may not be considered content. Even though they may convey important or sensitive information, they do not contain the message that you are actually trying to communicate to someone else, but rather information about the source or destination of that message.

"There's not a sharp difference between metadata and content. . . . It's more of a continuum."

Michael Morell, former CIA official

Electronic copy available at: https://ssrn.com/abstract=2573962

Although this distinction may appear clear, it quickly becomes blurry on closer examination. For example, technically speaking, a URL is very much a "delivery instruction;" it specifies the address of the web page that you are requesting. But it is also content: requesting a web page essentially means sending a message saying "please send me back the page found at this URL." In addition, a single URL reveals exactly which page was sought, and thus exactly what content was received. As a result, at least one court has noted that URLs "may be" content,[3] and the Department of Justice requires U.S. Attorneys to consult with an expert before obtaining URL information.[4]

Drawing a line between content and metadata is equally difficult with other forms of communications. As any high school senior awaiting college admissions decisions knows, the size of an envelope often says all you need to know; does that make "envelope size" content, and if so, does the same apply to the length of an email, or the types or sizes of any attachments? Is sending a friend request to a Facebook user content, and if so, does the same apply to a list of your friends? If the Web sites you visit or the text of the email you send and receive are content, what about keywords or advertising profiles generated directly from that site or email?[5] If your answers to an online survey are content, what about your inclusion on a list of individuals with a particular medical condition based on those answers?

In addition, whether information is content or metadata can depend not only on the type of information but also on the context in which it is created or used. This means that exactly the same information can be content in one situation and metadata in another. For example:

- Your location may or may not be content depending on context. If you call your friend and say "I am at Starbucks," the words you speak are content. If you use your smartphone to "check in" with Foursquare, that check-in is also content. But many courts have held that your cell carrier's record of the location of your phone at the exact same moment is not content.[6] And what if you take a picture or post a Tweet that you tag (intentionally or unintentionally) with your current GPS coordinates?

- The identity of your friends and contacts may or may not be content depending on context. If you write an email stating that "John is my friend," that statement is content. But it is less clear whether the fact that John is on your Facebook friend list is also content, even though it conveys exactly the same message.

In far too many situations, information may or may not be considered content depending on arbitrary interpretations or contextual factors that have little to do with the information itself. As Michael Morell, former CIA official and member of the President's Review Group on Communications Technologies, put it: "There's not a sharp difference between metadata and content. . . . It's more of a continuum."[7] But determining that information is not content can leave it with less privacy protection than content would receive. This remains true despite the ever-increasing value and sensitivity of metadata.

Electronic copy available at: https://ssrn.com/abstract=2573962

The Significance and Risks of Metadata

Even narrowly defined, metadata can expose sensitive information and present significant risks in the wrong hands. Some metadata, particularly associational information and location information, is inherently expressive, capable of directly exposing intimate details of an individual's life. In addition, the rise of "big data" has led to the development of tools that can take seemingly innocuous data points — drawn from consumer purchases, social networks, IP addresses, and more — and use them to ferret out hidden facts about a person. Finally, metadata can expose not only a person's activities but her identity as well. Both well-publicized corporate snafus and academic research demonstrate that it is often quite straightforward to take a supposedly anonymous collection of metadata and link it back to a specific person. In other words, metadata can reveal who we are, who we know, what we do and care about and plan to do next — essentially the same spectrum of sensitive information that could also be contained in the contents of a communication. In fact, metadata can even reveal things that we never intentionally communicated at all.

"Communicative" Metadata

Some forms of metadata are inherently communicative, directly revealing potentially intimate details about an individual without requiring any extra effort. For example, knowing with whom a person communicates — and when and how frequently — can expose a great deal of information, in some cases as much as or even more than the contents of those communications themselves. It's not hard to uncover an individual's dissatisfaction with his job, marital difficulties, or health status if you know that he is in frequent contact with a recruiter, divorce attorney, or cancer treatment center. And it's even easier to "infer" a person's sexual orientation or political allegiance if you can reveal her connection with the LGBT Choir or the local Tea Party chapter.

Location information also can communicate a great deal, presenting a direct risk to privacy and even physical safety. Knowing where an individual is can reveal whether he is "attending a religious service or a support meeting, visiting a doctor's office, shopping for an engagement ring, playing hooky from work, or spending an evening at the corner bar."[8] And because location information involves, well, location, the consequences can be far more than merely annoying; inadvertently-disclosed location information has been linked to numerous stalking cases.[9]

In many other situations, isolated pieces of metadata can be immediately revealing without the need for further analysis. The mere fact that a person purchased a pregnancy test, visited a web site focused on living with HIV, or wrote a check for thousands of dollars to a Las Vegas casino can clearly "tell a story" with potential repercussions. But because these pieces of data are not formally communications content, they may lack legal protection.

Electronic copy available at: https://ssrn.com/abstract=2573962

Aggregate Metadata and Data Mining

> **"It's much more intrusive than content. . . . You can see a call to a gynecologist, and then a call to an oncologist, and then a call to close family members."**
>
> **Susan Landau, Security Expert**

Although just one piece of metadata can provide a meaningful glimpse into a person's private life, aggregate metadata can reveal far more. Even small collections of metadata can expose a great deal; if someone "can see a call to a gynecologist, and then a call to an oncologist, and then a call to close family members," they likely have a very good guess as to what those calls were about.[10] By collecting and analyzing huge amounts of data, often of different types, companies and government actors can extract even more, including facts that individuals consciously choose not to reveal and even patterns that they may not be aware of. For example:

- Based on a woman's purchase history, Target can not only determine that she is likely to be pregnant even if she has not purchased any specific pregnancy-related items, but can even project the due date of her child. In one instance, the father of a teenage girl learned of his daughter's pregnancy because Target mailed coupons for pregnancy-related items to his home before his daughter told him.[11]

- Student researchers at MIT discovered that they could accurately determine the sexual orientation of an individual based solely on an analysis of his social network connections and their profiles.[12]

- Immersion, a tool developed by another set of students at MIT, uses basic email metadata (the addresses of the sender and recipient and the time the email was sent) to graphically illustrate how personal and professional networks evolve over time.[13] This kind of analysis can "remind you of former loves, illustrate the changing dynamics of your professional and personal networks over time, mark deaths and transitions in your life, and more."[14]

- Computer scientists found that a person's ethnicity and relationship status can be determined based solely on the person's own cell phone location information.[15]

- Analysts with the Office of the Privacy Commissioner of Canada determined that an IP address can be linked to posts on online forums and usernames on other sites and services, which can expose an individual's professional activities and interests, religious pursuits, health and personal issues, and more.[16] IP addresses can reveal information about the physical location of an Internet user as well.

In fact, mining metadata can not only expose sensitive information about the past, it can even allow an observer to predict future actions. For example, research has demonstrated that an individual's future location and activities can be predicted by looking for patterns in his friends and associates' location history.[17] A security expert also

Electronic copy available at: https://ssrn.com/abstract=2573962

warned that identifying phone calls from key executives at a company to or from a competitor, an attorney, or a brokerage can reveal the potential for a corporate takeover before any public announcement is made.[18] The possibility of not only exposing past actions but also future plans increases the risk of harm if metadata is inadequately protected.

Metadata and Identity

Finally, metadata can be used not only to reveal a person's activities but also her identity. In many cases, it takes surprisingly little metadata to identify an individual. For example, according to one researcher, "[t]he way we move . . . is so unique that four points [of location information] are enough to identify 95% of people."[19] Similarly, researchers have been able to take "anonymized" data sets from Netflix and AOL and re-identify many of the users whose data was released.[20] As a result, metadata privacy cannot be maintained simply by removing what is commonly known as personally identifiable information. In fact, one researcher has asserted that "[a]ny dataset that has enough information on people to be interesting to researchers also has enough information to be de-anonymized."[21]

Research shows that four points of location information are enough to identify 95% of people.

Metadata can expose a wide range of information about an individual, from his identity to his political beliefs, romantic entanglements, and health and financial concerns — but too often it lacks the privacy protections afforded to communication content. Law enforcement and national security agencies have stretched even these weak limits in their efforts to collect metadata, deploying surveillance programs that infringe on constitutional rights and have led to numerous incidents of abuse.[22]

PART II: Inadequate Protections for Metadata Lead to Abuse

There are several sources of privacy protection under U.S. law at both the federal and state level. Unfortunately, courts and lawmakers have frequently declined to extend these protections to metadata, often based on justifications that are increasingly inappropriate given the explosion of metadata and the tools to analyze it in the modern world.

Laws and Doctrines that Limit Protections for Metadata

Various laws in the United States, including federal and state constitutions and statutes, strive to protect the privacy of electronic communications, imposing strong standards including judicial oversight on any attempt to obtain the content of those communications. However, these same laws and court decisions frequently relegate

Electronic copy available at: https://ssrn.com/abstract=2573962

metadata to second-class status. In doing so, they leave the wealth of sensitive information exposed by metadata ripe for abuse.

The Fourth Amendment and the Third Party Doctrine

The Fourth Amendment to the U.S. Constitution protects individuals from unreasonable searches and seizures. In principle, it would seem that this protection would apply just as clearly to the capture and analysis of metadata as to the interception or disclosure of communications contents. However, courts have frequently applied a judicially-created exception to the Fourth Amendment, called the "third party doctrine," to hold that records of individuals' phone calls, location, Internet use, and more collected by companies lack constitutional protection.

In most instances involving data, the applicability of the Fourth Amendment turns on whether or not an individual has a "reasonable expectation of privacy." The reasonable expectation of privacy test was established in *Katz v. United States*, in which the Supreme Court held that an individual's phone calls were protected from warrantless interception even if they were made from a public telephone booth because he reasonably expected, and society as a whole accepted, that such calls would remain private.[23] However, in the 1970s a pair of pre-Internet Supreme Court cases held that an individual has no reasonable expectation of privacy (and thus no Fourth Amendment protection) for records held by a third party business.[24] This doctrine, known as the "third party doctrine" or "business records doctrine," frequently has been applied in cases involving metadata.

The third party doctrine has been widely challenged in recent years. In fact, two district courts have recently rejected the third party doctrine as it applies to metadata,[25] and U.S. Supreme Court Justice Sonia Sotomayor has explicitly called for it to be reevaluated.[26] However, to date many courts continue to apply the third party doctrine and state that metadata — including location records, information about emails, and more — held by a third party is excluded from the protections of the Fourth Amendment.[27]

The Electronic Communications Privacy Act

The Electronic Communications Privacy Act (ECPA) was passed in 1986 as a direct response to Congressional concern that the third party doctrine would deprive Americans of their right to privacy in the emerging digital world. ECPA protects non-content as well as content information; however, it explicitly differentiates between the two and provides lesser protection to metadata than to content.

Under ECPA, a search warrant is required to compel disclosure of the contents of a communication that is in "electronic storage" for no more than 180 days.[28] However, ECPA allows law enforcement agents to obtain metadata without satisfying the requirements of a search warrant. Under ECPA, a governmental entity need only provide "specific and articulable facts showing that there are reasonable grounds to believe that . . . records or other information sought[] are relevant and material to an ongoing criminal investigation" in order to obtain

Electronic copy available at: https://ssrn.com/abstract=2573962

"information pertaining to a subscriber to or customer of [an online] service (not including the contents of communications)."[29] ECPA also permits demands for specific kinds of information, including a subscriber's device or account address, without a court order, requiring only a subpoena to compel disclosure. As a result, state and local police as well as federal law enforcement have broad authority to obtain metadata without a warrant.

The Foreign Intelligence Surveillance Act and the Patriot Act

While ECPA provides the framework for access to electronic information, including metadata, for law enforcement, a separate process established by the Foreign Intelligence Surveillance Act allows the National Security Agency to issue its own demands for information. Paralleling the lower Fourth Amendment requirements for searches and seizures in the law enforcement context, FISA is generally more permissive than ECPA in allowing access to information.

In particular, the USA PATRIOT Act of 2001 amended FISA to give the NSA broad latitude to compel the disclosure of "business records" in order to "obtain foreign intelligence information" or "protect against international terrorism or clandestine intelligence activities."[30] While FISA requires that an order compelling the disclosure of business records be issued by a judge, the threshold for obtaining such an order is much easier to meet than that required to obtain a search warrant. As a result, it appears that nearly 100 percent of applications for orders to compel business records are granted even though a single order can target an entire database of records.[31] The "call record database" exposed by Edward Snowden was authorized under Section 215 of the Patriot Act.

The Patriot Act also greatly expanded the FBI's authority to issue National Security Letters.[32] These letters allow the FBI, without judicial oversight, to compel service providers to turn over certain metadata about subscribers. They also routinely prevent the service from notifying the subscriber about the Letter, even if the FBI does not pursue any further investigation.

Failing Justifications for the Content/Metadata Distinction

Metadata is capable of conveying or revealing information that is in many cases just as sensitive as the contents of a communication, yet courts and lawmakers have frequently afforded it less protection. The distinction between content and metadata has been justified with various theories and doctrines. In the end, however, all of these justifications suffer the same fatal flaw: an outdated assumption that individuals do not expect their metadata to be protected. This assumption is no longer valid in the modern data-driven world.

The Envelope Metaphor: "Non-Content Is Not Private"

One of the most commonly-used metaphors to illustrate the difference between content and non-content is a letter enclosed in an envelope. The inside of the envelope holds its contents; those contents are sealed away from prying

Electronic copy available at: https://ssrn.com/abstract=2573962

eyes and thus private. The outside of the envelope, in contrast, is public; anything written there is voluntarily made visible for the world to see.

The first problem with the envelope metaphor, and the general assumption that metadata is voluntarily exposed, is that a great deal of metadata related to electronic devices and communications is created without the individual's actual knowledge, let alone consent. For example, simply carrying around a powered-on cell phone creates an ongoing record of a person's location (because a cell phone continuously interacts with nearby relay towers in order to remain connected to the network). However, few users are aware of the fact that their phone is constantly generating information about their location. And although one New York magistrate judge recently issued an opinion "straight from the Twilight Zone" instructing users to turn off their cell phones if they want privacy, the vast majority of cell phone users do expect privacy in their location information even with their phone turned on.[33]

Moreover, even users who are aware of the kinds of metadata collected and exposed by electronic communications services have little choice as to whether to share that information. A letter can be sent without a return address and dropped off in a mailbox, greatly reducing the possibility of connecting the sender to the recipient. It is much more difficult, and in many cases impossible, to voluntarily decline to share information about electronic communications. Individuals cannot make or receive a cell phone call without having that information tracked by their carrier. Internet communications require the use of IP addresses that can be logged. Thus, there is little "voluntary" surrender of privacy in an electronic communication analogous to the readily-available choices available to a letter writer.

Finally, in many cases so-called "envelope data" is not in fact exposed to observers. An increasing number of Internet sites and services use encryption to ensure that other parties are prevented from viewing not only the "content" of an interaction (such as the photo that a user is uploading to a social network) but anything associated with that content (such as the photo's title, location, or timestamp). In other words, they ensure that all of their users' metadata are in fact "inside the envelope."

The Third Party Doctrine: "Non-Content Is Not Yours"

A second justification offered for lesser protections for metadata derives from the fact that metadata often is not consciously created by an individual. Instead, it may be generated by a third party, such as an online service or cell carrier. Following this logic, courts have created and utilized the third party doctrine to hold that records held by a third party fall outside of the protections established by the Fourth Amendment. If these records belong to the third party rather than to the individual, the argument goes, how can the individual reasonably expect them to be kept private?

Electronic copy available at: https://ssrn.com/abstract=2573962

This justification is directly contradicted by the actions of both users and companies demonstrating that users do in fact expect such information to be kept out of the government's hands. Many companies recognize the importance of building user trust by safeguarding the data entrusted to them. Companies increasingly use encryption and security techniques to protect the data that they receive — in some cases by creating systems that prevent the company itself from accessing user data at all.[34] In addition, several companies have stated that they require a search warrant for much of the data they retain regardless of the third party doctrine.[35] Efforts to protect user data have increased since the NSA surveillance programs were revealed.

The adoption of such policies, and the attitude of the companies towards user data, undermines the argument that users neither expect nor deserve privacy protection for their metadata simply because that data is in the hands of a third party. Instead, these third parties stand beside their users, supporting and asserting their users' rights to protect their personal information from warrantless surveillance. And such a stance further supports the contention that users are acting reasonably in expecting their data to be protected. According to a district court judge in Washington, D.C., the evolving relationship between users and online services has led not to a diminished expectation of privacy but instead to "a greater expectation of privacy and a recognition that society views that expectation as reasonable."[36]

"Non-Content Is Non-Sensitive"

Finally, government officials have asserted repeatedly that metadata is simply not sensitive. In the context of the recent NSA scandal, President Obama was quick to reassure the American people that "nobody is listening to your phone calls."[37] Nor is he alone in asserting that information about communications and activities is inherently less sensitive than the actual words used to communicate.

However, this assertion is in direct conflict with the ever-increasing focus on the collection and analysis of "behavioral data" and other metadata about individuals. Metadata can reveal just as much intimate information about an individual as the contents of her communication, including personal information about health, sexuality, or relationships, or other sensitive information such as ongoing discussions about a corporate takeover[38] or the sources that reporters use to break news about government activities.[39]

As technology advances into new frontiers of our lives, it is increasingly difficult to imagine any information that could be communicated as content but somehow impossible to obtain via metadata. Wearable devices that measure physiological signals can potentially expose an individual's emotional state.[40] Patterns of communications can be used to predict where an individual is likely to go in the future.[41] There is simply no justification for asserting that metadata is categorically less sensitive and thus less deserving of protection than content in the modern world.

Electronic copy available at: https://ssrn.com/abstract=2573962

Consequences of Lesser Protection: Mass Surveillance and Frequent Abuses

Given the value of metadata and the lower threshold require to obtain it than to obtain content, it comes as no surprise that law enforcement and national security entities frequently demand metadata from third parties. Even so, the actual scope of these demands is astonishing, with law enforcement accessing millions of records and tracking individuals for months at a time without a warrant and the NSA attempting to collect information about every single phone call and Internet communication. This has resulted in uses of data, both within and outside of the scope of agency policies, that intrude deeply on the privacy rights of Americans.

NSA Collects Records of Every Single Phone Call Carried by Major Providers

By far the most graphic example of the abuse of metadata's weak protections is the still-emerging evidence of massive collection of communications metadata by the National Security Agency. Documents released by whistleblower Edward Snowden reveal that the NSA has continuously issued orders demanding records about the parties to and duration of every single phone call carried by major U.S. telecommunication providers.[42] In addition, the NSA has apparently collected detailed location information about both Americans and foreign nationals from companies around the world.[43] A separate NSA program, since terminated, was designed to vacuum up vast amounts of metadata about Internet activity.[44]

There are several different factors that have allowed these programs to expand from targeted observation of known persons of interest to massive surveillance of the entire nation. One is the fact that the Foreign Intelligence Surveillance Court has accepted the NSA's extremely broad interpretation of "relevance" and authorized demands for massive datasets — such as records about every call carried by Verizon's network over a 3 month period — as long as the NSA claimed that some piece of information in the set had foreign intelligence value.[45] But the underlying cause is the lack of legal protection for metadata. In evaluating the recent Fifth Circuit decision reaffirming the lack of Fourth Amendment protection for metadata, a former Justice Department attorney noted that "if the Fifth Circuit had gone the other way, [the NSA] would not be able to [obtain phone call location data] pursuant to section 215."[46] One district court, in Washington, D.C., recently held that the NSA's metadata collection program likely violates the Fourth Amendment — but the court's order prohibiting continued metadata collection has been stayed pending an appeal.[47] In the meantime, the massive collection of metadata appears to be continuing.[48]

Abuses of this massive dataset have already been uncovered. A recent article in the Washington Post revealed that the NSA's internal privacy rules (to say nothing of constitutional limits) were violated "thousands of times per year."[49] Recent reports suggest that the NSA has provided the DEA with "tips" based on its bulk surveillance data despite express prohibitions on such data sharing.[50] The NSA has even come up with its own internal code,

Electronic copy available at: https://ssrn.com/abstract=2573962

LOVEINT, to describe incidents where an analyst abused his or her authority in order to monitor a romantic interest.[51] Given the vast scope of these programs, it seems safe to assume that even more violations remain undiscovered.

Department of Justice Seizes AP Phone Records

Just prior to the revelations concerning the NSA's massive surveillance program, the Associated Press (AP) learned that the Department of Justice had secretly obtained two months' worth of calling records for 20 AP offices and journalists.[52] The seizure, conducted in the wake of a series of leaks concerning the Central Intelligence Agency, potentially exposed the identities of every confidential source working with the Associated Press on any issue, as well as insight into the AP's activities and operations that, as the AP asserted, "the government has no conceivable right to know." And while the scope of the intrusion is unsurprising in light of more recent revelations about the NSA, the fact that a journalistic endeavor was the specific target of the investigation raises serious concerns about press freedom and freedom of expression.

"There can be no possible justification for such an overbroad collection of the telephone communications of The Associated Press and its reporters. These records potentially reveal communications with confidential sources across all of the news gathering activities undertaken by The A.P. during a two-month period [and] operations that the government has no conceivable right to know."

Gary Pruitt, President & CEO,
The Associated Press

Law Enforcement Collection of Location, Cell, and Online Data

While most law enforcement agencies lack both the authority and the capacity to monitor every communication in the country, they are nonetheless able to take advantage of lax protections for metadata to acquire vast amounts of information about individuals.

Location information has been a particular target of warrantless demands. According to Sprint's own representative, law enforcement agents "pinged" cell phones on the company's network for location information more than 8 million times over a 13 month period.[53] A wide-ranging survey by the ACLU uncovered a variety of practices law enforcement use to obtain location information; while many law enforcement agencies routinely obtain a warrant prior to demanding location information, a larger number obtain such information without a warrant.[54]

Law enforcement regularly seeks other forms of information as well. Several companies now publish "Transparency Reports" documenting the number of demands for information about individuals that they receive.[55] These reports invariably demonstrate an increasing number of demands for information, including demands made via legal process other than a search warrant. Letters sent by Senator Edward Markey (D-MA) revealed that

Electronic copy available at: https://ssrn.com/abstract=2573962

telecommunications carriers received over 1 million demands for information (including both communications contents and metadata) in 2012.[56] And Freedom of Information Act requests exposed the "Hemisphere" program in which AT&T employees providing Drug Enforcement Agency officers with warrantless access to cell phone records dating back to 1987.[57]

In addition, law enforcement agencies increasingly collect information directly, including location information and online data. Police across the country are using automated license plate recognition systems (ALPRs) to fill databases with millions of records, including information about political protesters.[58] Devices called "stingrays" have apparently been used to directly locate cell phones without carrier involvement.[59] And many agencies are looking to "mine" information from social networks and elsewhere on the Internet to create robust profiles of individuals.[60]

In various instances, law enforcement has collected metadata for purposes that are clearly inappropriate. For example:

- In 2010, Michigan police sought information about every single phone located near the site of a planned labor protest without a warrant.[61]
- A Tennessee sheriff requested the location of his daughter when she was out past her curfew.[62]
- A police chief in South Carolina obtained four "tower dumps" providing information about every cell phone within range of two separate cell towers after his personal vehicle was burglarized.[63]

Given the rise in demands for metadata and the fact that few demands for such information ever see the light of day, this is likely only the tip of the iceberg. The only way to prevent ongoing misuse and abuse of this information is to ensure that the privacy protections for metadata — including both limits on access to metadata and transparency and oversight when access is permitted — match those for content.

PART III: Efforts to Enhance Protections for Metadata

Although metadata's protections have generally failed to keep pace with its growing value and sensitivity, courts and legislatures have begun to erect meaningful protections for at least some categories of metadata. Metadata related to the freedoms of expression and association in the First Amendment to the U.S. Constitution and various state constitutions receives protection that often matches, and in some cases exceeds, the protections traditionally afforded to content. More recently, state legislatures have begun to recognize the inherent sensitivity of location

Electronic copy available at: https://ssrn.com/abstract=2573962

information and provide protection for such metadata. This represents a critical step in the right direction, away from the distinction between content and metadata and towards a privacy regime that protects data based on its sensitivity and potential for harm. The recent Supreme Court case *United States v. Jones* and its progeny provide new hope for a second critical step: revising or outright rejecting the third party doctrine and clarifying that Fourth Amendment protections apply to information, including metadata, even if it is held by a third party.

Legal Protections for Expressive and Associational Information

One type of metadata that receives significant legal protection is data concerning expressive and associational activities. The First Amendment to the U.S. Constitution, enshrining the rights to freedom of expression and association, and its state constitutional counterparts provide both the guiding principle and the legal underpinning for many of the protections for expressive and associational data. State and federal lawmakers have expanded upon these protections to ensure that individuals can freely distribute and receive content and form groups and associations without the chilling effects of constant exposure of their personal information.

Anonymity and Identifying Metadata

The Supreme Court has long held that "[p]rotections for anonymous speech are vital to democratic discourse. Allowing dissenters to shield their identities frees them to express critical minority views."[64] As a result, the Court has consistently upheld the right to anonymous speech, striking down laws prohibiting anonymous leafleting and clearly asserting the value of anonymity in the modern world.[65]

Numerous courts, including the Supreme Court, have recognized that the right to speak anonymously on the Internet can limit the compelled disclosure of metadata identifying Internet users.[66] In the civil context, there is a consensus among the courts that the fundamental idea of First Amendment protection of anonymous speech must be balanced against the power of a civil litigant to obtain identifying information.[67] In addition, companies such as Twitter have successfully challenged law enforcement efforts to unmask anonymous online speakers.[68]

> **"Once the government can demand of a publisher the names of the purchasers of his publications . . . [f]ear of criticism goes with every person into the bookstall . . . [and] inquiry will be discouraged."**
>
> ***United States v. Rumely***

Expressive Metadata

Numerous state and federal laws also protect the privacy of metadata that is related to expressive activity. These laws recognize that such metadata may both reveal sensitive information about an individual and implicate the freedoms guaranteed by the First Amendment.

Electronic copy available at: https://ssrn.com/abstract=2573962

Decisions by federal and state courts have consistently extended constitutional protections to expressive metadata. In *United States v. Rumely*, the Supreme Court found it unconstitutional for a bookseller to be convicted for refusing to provide the government with a list of individuals who had purchased political books.[69] Similarly, a D.C. district court barred the government from compelling a local bookstore to disclose the books purchased by Monica Lewinsky, holding that the First Amendment required the government to "demonstrate a compelling interest in the information sought . . . [and] a sufficient connection between the information sought and the grand jury investigation . . . ."[70] The Colorado Supreme Court similarly held that book records were clearly protected under the free speech provision of the Colorado state constitution and subject to a heightened standard.[71] And in 2010, a federal court in the state of Washington quashed a demand for records of Amazon customer purchases, holding that "[t]he First Amendment protects a buyer from having the expressive content of her purchase of books, music, and audiovisual materials disclosed to the government. Citizens are entitled to receive information and ideas through books, films, and other expressive materials anonymously."[72]

"The First Amendment protects a buyer from having the expressive content of her purchase of books, music, and audiovisual materials disclosed to the government. Citizens are entitled to receive information and ideas through books, films, and other expressive materials anonymously."

Amazon. com LLC v. Lay

Lawmakers have also pursued legislation in order to protect expressive metadata. The federal Video Privacy Protection Act and Cable Act place restrictions on demands for records of video rentals and cable records respectively.[73] The California Reader Privacy Act, enacted in 2011, requires a showing that there is probable cause to believe that evidence of a crime will be revealed, there is a compelling need to obtain such evidence, and there is no less intrusive means of doing so, prior to authorizing the compelled disclosure of records held by physical or online book providers.[74] And nearly every state in the country has passed laws ensuring the privacy of library records.[75]

Associational Metadata

Finally, the First Amendment also protects the privacy of associational information. In a seminal case, *NAACP v. Alabama*, the Supreme Court held that constitutional protections for freedom of expression and freedom of association include a right to privacy with regard to associational information.[76] Thus, individuals and groups alike have the right to protect their expressive abilities by keeping records of their members and activities private. Two recent court cases, *ACLU v. Clapper* and *First Unitarian Church of Los Angeles v. NSA*, have expressly challenged the NSA's metadata collection program on these grounds.[77]

Legal Protections for Location Information

Not only does the improper use or disclosure of location information present a risk of exposing a wide range of

Electronic copy available at: https://ssrn.com/abstract=2573962

highly sensitive information about an individual, revealing an individual's current location can present a wider range of threats including burglary or even physical assault. In response, courts and state legislatures have in many cases established meaningful protections for location information.

Most significantly, the Supreme Court recently affirmed a decision holding that long-term warrantless tracking of a vehicle was unconstitutional under the Fourth Amendment. Although the majority opinion in that case relied on a trespass-based theory, four Justices would have held that tracking a vehicle without a warrant for an extended period of time required a warrant even without the physical invasion of placing a tracking device on the vehicle.[78] A fifth Justice, Justice Sotomayor, not only suggested that she was sympathetic to that view but expressly called into question the validity of the third party doctrine in the modern age.[79]

State courts have also addressed the specific question of location surveillance under their state constitutions. The Supreme Court of New Jersey has explicitly applied a constitutional warrant requirement to law enforcement demands for location information.[80] The highest court of Massachusetts held that surveillance by means of a GPS tracking device is a violation of the state constitutional right to privacy, but has not (yet) extended that decision to all forms of location information.[81]

In addition, state and federal legislators have attempted to address the issue of location privacy. In May of 2013, Montana became the first state to require law enforcement to obtain a search warrant prior to obtaining any form of location information,[82] followed by Maine later in the year.[83] Several other states are considering or have considered similar bills.[84] There have also been several bills introduced in Congress to safeguard the privacy of location information.[85]

Challenges to Warrantless Metadata Collection

In the aftermath of the Snowden revelations, there have been various responses to the NSA bulk metadata programs in the courts, Congress, and even the administration. In *Klayman v. Obama*, the District of Columbia District Court held that the metadata program was unconstitutional, a decision that has impacted the discussion of the program even while it is being appealed.[86] Other direct challenges to the program, including a lawsuit filed by the ACLU, have been filed as well.[87] In Congress, the USA Freedom Act was introduced by Senator Patrick Leahy (D-VT) and Representative Jim Sensenbrenner (R-WI) in response to revelations about the NSA's bulk metadata program. The proposed bill would prevent the bulk collection of metadata, allowing data only to be collected about suspected terrorists or persons

"The government surveillance programs conducted under the Foreign Surveillance Intelligence Act are far broader than the American people previously understood. It is time for serious and meaningful reforms so we can restore confidence in our intelligence community."

Rep. Jim Sensenbrenner (R-WI)

Electronic copy available at: https://ssrn.com/abstract=2573962

directly in contact with suspected terrorists.[88] In January 2014, President Obama announced his intention to examine alternatives to the bulk collection of metadata and to require court authorization for access to the current database,[89] implementing in part the recommendations of the President's Review Group on Intelligence and Communications Technologies.[90] The Review Group's report specifically called out the issue of metadata, recommending that the government "commission a study of the legal and policy options for assessing the distinction between meta-data and other types of information" including perspectives focused on privacy and civil liberties.[91] Finally, later that same month the Privacy and Civil Liberties Oversight Board issued a report questioning both the efficacy and the legality of the NSA's telephone metadata program and calling for its cancellation.[92]

In addition to responses specific to the NSA metadata program, efforts to revise or overturn the third party doctrine appear to be gaining momentum. The prospects for meaningful reform have been boosted considerably by Justice Sotomayor's concurrence in *United States v. Jones*, where she expressly stated her willingness to reconsider the third party doctrine and described it as "ill-suited to the digital age, in which people reveal a great deal of information about themselves to third parties in the course of carrying out mundane tasks."[93] Some federal courts have in fact narrowed or outright rejected the third party doctrine. In *United States v. Warshak,* the Sixth Circuit held that the third party doctrine did not apply to online emails. In *United States v. Powell*, a court in the Eastern District of Michigan held that the Fourth Amendment required a warrant for extended tracking of an individual's location even without a physical trespass.[94] And in *Klayman v. Obama* a District of Columbia district court judge held specifically that the third party doctrine was inapplicable in finding that the NSA's metadata collection was unconstitutional.[95]

Several states, including California, have explicitly rejected the third party doctrine in state law, holding that their constitutional privacy protections apply even when information is held by a third party.[96] These holdings not only provide protections for citizens of those specific states, but also help to inform the national debate and provide strong evidence that individuals do in fact reasonably expect that their information — including metadata — will remain private even if held by a third party.

PART IV: Establishing Privacy Protections for Metadata

Emerging legal protections for expressive and associational records and location information demonstrate that courts and lawmakers recognize the need to look beyond the simple content/metadata distinction in certain instances. However, the resulting patchwork regime is still likely to leave many types of metadata without adequate

Electronic copy available at: https://ssrn.com/abstract=2573962

protection even if they expose exactly the same sensitive information as better-protected forms of data. Moreover, efforts to maintain such a patchwork solution will inevitably be outpaced by new forms of metadata arising from the rapid evolution of technology. Instead, we need an approach that is forward-looking and comprehensive: completely abandoning the distinction between metadata and content and replacing it with robust privacy protections for all sensitive data that better matches the modern digital world.

There are several principles that can help guide the development of such protections:

Protect Sensitive Information Regardless of Form

In order to adequately protect individual privacy, legal protections must apply to all sensitive personal information, regardless of the type or category of that information. This is the only way to produce a forward-looking regime that is capable of keeping pace with the rapid evolution of technology.

One example of a privacy framework that rejects the content/metadata distinction is found in California. The right to privacy in the California Constitution, which was drafted expressly to address concerns about informational privacy, applies to all "personal information" (information that can be linked to a specific individual), not merely content. By basing privacy protections on information that is connected to a specific person rather than on the form of the information, this provision supports a much more robust privacy regime than one based on the content/metadata paradigm.

Protect Sensitive Information Regardless of Possessor or Storage Location

The idea of robust privacy protections for metadata is fundamentally inconsistent with the third party doctrine. While there are various types of metadata that individuals generate and retain on their own device or otherwise in their possession, the overwhelming majority of metadata is created or captured by third parties. Several states, including California, have already rejected the third party doctrine under their own state constitution. The federal judiciary should take up the suggestion offered by Justice Sotomayor and do the same.

Protect Sensitive Information Derived from Data Aggregation

Comprehensive protection of metadata must also take into account the fact that large sets of data can reveal sensitive information that cannot be inferred from any specific element in that set. This means that privacy protections need to apply not only to data directly collected from an individual but also to any inferences or derivative information generated through the analysis of that data. And it requires carefully evaluating the circumstances where law enforcement can compile data about individuals without any sort of legal process.

Electronic copy available at: https://ssrn.com/abstract=2573962

One approach to this problem is the "mosaic theory" of privacy presented by the D.C. Court of Appeals in *United States v. Maynard*. The essential holding of *Maynard* is that aggregating bits of otherwise-unprotected information into a "mosaic" that can reveal far more information than the bits in isolation should trigger privacy protections.[97] Thus, while a single observation of an individual in public may not be subject to Fourth Amendment protection, focused surveillance of the individual through technological means does trigger that protection.

There has been extensive debate about the practicality of the mosaic theory. Some scholars have noted the potential hindrance to law enforcement, arguing that officers have no way of knowing when they have crossed an invisible threshold from the legitimate acquisition of unprotected information without legal process to a "search" requiring such process.[98] Others have expressed confidence in the ability of the courts to turn abstract concepts into concrete rules, noting that this is precisely how current Fourth Amendment doctrines ranging from the Miranda rules to requirements based on the "reasonable expectation of privacy" standard have evolved.

While potentially imperfect, the mosaic theory is still clearly preferable to the status quo. Like the mosaic theory, current law has its own set of boundaries and dividing lines (such as the "persnickety, but persistent question"[99] about what is and is not content) that have proven difficult to clearly define. There is no reason to believe courts will be any less successful in creating practical rules concerning the mosaic theory than they have been under current doctrine. More importantly, existing law is based on the content/metadata paradigm that is already outdated and will fall further and further behind the realities of modern technology. The mosaic theory, on the other hand, offers the potential to effectively evolve along with our society.

Provide Tools and Guidance for Law Enforcement Access to Metadata

Comprehensive privacy protections for metadata are not feasible if the protections they provide render effective enforcement of the law impossible. Thus, such efforts must meet two criteria. First, they must provide law enforcement with the tools necessary to protect the public. Second, they must provide law enforcement with the guidance necessary to comply with the requirements of the law.

Lawmakers and courts should provide law enforcement with the authority needed to carry out its mission, including identifying potential suspects or witnesses in an investigation, while ensuring that this authority does not enable extensive metadata collection without a warrant. Courts should be authorized to provide access to narrowly-defined categories of information such as ECPA's "subscriber information" without a warrant as long as there are well-defined limits preventing the information from being used for other purposes or simply stockpiled. In addition, any request for identifying information should be carefully examined to ensure that the law enforcement interest outweighs any First Amendment consideration for anonymous speech.

Electronic copy available at: https://ssrn.com/abstract=2573962

In addition, any rules that limit access to metadata must be accompanied by guidance allowing the involved parties to follow the rules and making it clear when a search warrant is required. Fortunately, legislators and courts have proven adept at crafting rules that are administrable and addressing problems as they arise. There is no reason that they should be unable to do so again while creating rules that protect the privacy of all sensitive information, including metadata.

Ensure Appropriate Transparency and Oversight

Finally, as the NSA fiasco has illustrated, privacy law's impact will be muted if government agencies are allowed to conduct surveillance in secret based on their own interpretation of their authority. As a result, the President's Review Group on Intelligence and Communications Technology has specifically endorsed greater transparency about any ongoing surveillance programs, including metadata aggregation and analysis.[100] Any future data privacy regime must inform the public as much as possible about any planned or ongoing efforts to collect information about individuals. This should include both mandatory government reporting akin to the current Wiretap Report[101] and explicit permission for companies to produce independent transparency reports encompassing the broadest possible set of demands for user information. Robust transparency for both the national security and law enforcement sectors will help to ensure compliance with current law. It also will help privacy and security co-exist and evolve together as new forms of metadata and new techniques for collecting and analyzing it emerge.

CONCLUSION

In the modern world, metadata speaks loudly. It can expose an individual's sexual orientation, physical or mental health issues, political or social activities, and more. Unfortunately, even though metadata can reveal just as much about an individual as the contents of a communication, it is nonetheless treated as less important by many of our courts and laws. As a result, the collection of metadata has allowed both mass surveillance of the U.S. population on an unprecedented scale and abuses of this information by those in positions of authority.

There is no longer any justification for treating metadata as inherently less sensitive and less deserving of privacy protection than communication contents. Instead, we need a modern privacy regime that protects information based on its sensitivity, not its form. As the President's committee report noted, this requires rethinking the role of metadata in society and reexamining — and ultimately rejecting — legal rules based on the outdated distinction between content and metadata.

Electronic copy available at: https://ssrn.com/abstract=2573962

ENDNOTES

[1] Optiver Australia Pty. Ltd. & Anor. V. Tibra Trading Pty. Ltd. & Ors., No. C 12-80242 EJD (PSG), 2013 WL 256771 (N.D. Cal. Jan. 23, 2013) at *1.
[2] *See Metadata*, http://dictionary.reference.com/browse/metadata.
[3] United States v. Forrester, 512 F.3d 500, 510 n.6 (9th Cir. 2007).
[4] U.S. Dep't of Justice, U.S. Attorney's Manual, Title 9-7.500.
[5] *Cf.* Optiver, 2013 WL 256771 at *2 (holding that "content" under the Stored Communications Act includes any and all information about messages specifically containing a designated keyword). If obtaining metadata about a set of messages identified by a content-based keyword is equivalent to obtaining content, the same might apply to profiling information generated from content.
[6] *See, e.g.*, In Re Application of the United States of America for Historical Cell Site Data, No. 11-20884 (5th Cir. July 30, 2013), *available at* http://www.volokh.com/wp-content/uploads/2013/07/11-20884_Documents.pdf; United States v. Graham, 11-0094, LEXIS 26954 (D. Md. Mar. 1, 2012).
[7] Julian Sanchez, *Obama Backs Off Real NSA Reform*, DAILY BEAST, Jan. 15, 2014, http://www.thedailybeast.com/articles/2014/01/15/obama-backs-off-real-nsa-reform.html.
[8] ACLU of Northern California, *Location-Based Services: Time for a Privacy Check-In* 5, http://aclunc-tech.org/files/lbs-privacy-checkin.pdf.
[9] More than one in four stalking victims interviewed in a recent study reported some form of cyberstalking was used. Of these, GPS technology was involved in 10 percent of the electronic monitoring of stalking victims. U.S. Dep't. of Justice Special Report, *Stalking Victimization in the United States* at 5 (Jan. 13, 2009), *available at* http://www.ovw.usdoj.gov/docs/stalking-victimization.pdf.
[10] Jane Mayer, *What's the Matter with Metadata?*, NEW YORKER, June 6, 2013, http://www.newyorker.com/online/blogs/newsdesk/2013/06/verizon-nsa-metadata-surveillance-problem.html.
[11] Kashmir Hill, *How Target Figured Out a Teen Girl Was Pregnant Before Her Father Did*, FORBES, Feb. 16, 2012, http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/.
[12] Carolyn Y. Johnson, *Project 'Gaydar'*, BOSTON.COM, Sep. 20, 2009, http://www.boston.com/bostonglobe/ideas/articles/2009/09/20/project_gaydar_an_mit_experiment_raises_new_questions_about_online_privacy/.
[13] Immersion, https://immersion.media.mit.edu/.
[14] Kade Crockford, *Graphs by MIT Students Show the Enormously Intrusive Nature of Metadata*, FREE FUTURE, Jan. 7, 2014, https://www.aclu.org/blog/technology-and-liberty-national-security/graphs-mit-students-show-enormously-intrusive-nature.
[15] Yaniv Altshuler et al., *Incremental Learning with Accuracy Prediction of Social and Individual Properties from Mobile-Phone Data*, *in* PRIVACY, SECURITY, RISK AND TRUST (PASSAT), 2012 INTERNATIONAL CONFERENCE ON SOCIAL COMPUTING (SOCIALCOM), *available at* http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6403618.
[16] *See* OFFICE OF THE PRIVACY COMMISSIONER OF CANADA, WHAT AN IP ADDRESS CAN REVEAL ABOUT YOU, May 2013, http://www.priv.gc.ca/information/research-recherche/2013/ip_201305_e.asp.
[17] *See, e.g.*, E. Cho et al., *Friendship and Mobility: User Movement in Location-Based Social Networks*, *in* PROC. 17TH ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY AND DATA MINING 1082–90 (2011).
[18] Jane Mayer, *What's the Matter with Metadata?*, NEW YORKER, June 6, 2013, http://www.newyorker.com/online/blogs/newsdesk/2013/06/verizon-nsa-metadata-surveillance-problem.html.
[19] Jason Palmer, *Mobile Location Data 'Present Anonymity Risk,'* BBC.COM, Mar. 25, 2013, http://www.bbc.co.uk/news/science-environment-21923360.
[20] *See* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failures of Anonymization*, 57 UCLA L. REV. 1701 (2010), http://www.uclalawreview.org/?p=1353.
[21] Pete Warden, *Why You Can't Really Anonymize Your Data*, O'REILLY STRATA, May 17, 2011, http://strata.oreilly.com/2011/05/anonymize-data-limits.html.
[22] Privacy concerns related to the use of metadata by non-governmental entities such as corporations are outside the scope of this paper.
[23] Katz v. United States, 389 U.S. 347 (1967).
[24] Smith v. Maryland, 442 U.S. 735 (1979); United States v. Miller, 425 U.S. 435 (1976).
[25] *See* Klayman v. Obama, 13-cv-00851 (D.D.C. Dec. 16, 2013), *available at* http://legaltimes.typepad.com/files/obamansa.pdf, *and* United States v. Powell, _ F.Supp.2d _, 2013 WL 1876761 (E.D. Mich May 3, 2013).
[26] United States v. Jones, 565 US _, 132 S.Ct. 945, 957 (2012) (Sotomayor, J., concurring).
[27] *See, e.g.*, American Civil Liberties Union v. Clapper, 13 Civ. 3994 (WHP) (S.D. N.Y. Dec, 27, 2013) (memorandum & order to dismiss); In Re Application of the United States of America for Historical Cell Site Data, No. 11-20884 (5th Cir. July 30, 2013), *available at* http://www.volokh.com/wp-content/uploads/2013/07/11-20884_Documents.pdf; United States v. Graham, 11-0094, LEXIS 26954 (D. Md. Mar. 1, 2012).
[28] 18 U.S.C. § 2703(a) (2012).
[29] *Id.* § 2703(d).
[30] 50 U.S.C. § 1861(a)(1) (2012).
[31] *See* Erika Eichelberger, *FISA Court Has Rejected .03 Percent of Surveillance Requests*, MOTHER JONES, June 10, 2013, http://www.motherjones.com/mojo/2013/06/fisa-court-nsa-spying-opinion-reject-request.
[32] 18 U.S.C. § 2709 (2012).
[33] Chris Soghoian, *Federal Judge: Only Powered-Off Cell Phones Deserve Privacy Protections*, FREE FUTURE, http://www.aclu.org/blog/technology-and-liberty-national-security/federal-judge-only-powered-cell-phones-deserve-privacy.
[34] For example, Spideroak is a "zero-knowledge" online file sharing and cloud backup service that has no access to user data. Spideroak, https://spideroak.com.
[35] *See* Dieter Bohn, *Google, Microsoft, Yahoo!, Facebook Say They Require Warrants to Give Over Private Content*, The Verge, Jan. 26, 2013, http://www.theverge.com/2013/1/26/3917684/google-microsoft-yahoo-facebook-require-warrants-private-content.

[36] Klayman v. Obama, 13-cv-00851 (D.D.C. Dec. 16, 2013).
[37] Matthew DeLuca, *Obama: 'Nobody Is Listening to Your Phone Calls,'* NBC NEWS, June 7, 2013, http://usnews.nbcnews.com/_news/2013/06/07/18824941-obama-nobody-is-listening-to-your-telephone-calls.
[38] Jane Mayer, *What's the Matter with Metadata?*, NEW YORKER, June 6, 2013, http://www.newyorker.com/online/blogs/newsdesk/2013/06/verizon-nsa-metadata-surveillance-problem.html.
[39] *See* Ravi Somaiya, *Head of the A.P. Criticizes Seizure of Phone Records*, N.Y. TIMES, May 19, 2013, http://www.nytimes.com/2013/05/20/business/media/head-of-the-ap-criticizes-seizure-of-phone-records.html
[40] *Quantified Self Guide to Self-Tracking* Tools, http://quantifiedself.com/guide/.
[41] *See* Cho, *supra* note 17.
[42] Glenn Greenwald, *NSA Collecting Phone Records of Millions of Verizon Customers Daily*, GUARDIAN (UK), June 5, 2013, http://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order.
[43] Barton Gellman and Ashkan Soltani, *NSA Tracking Cellphone Locations Worldwide, Snowden Documents Show*, WA. POST, Dec. 4, 2013, http://www.washingtonpost.com/world/national-security/nsa-tracking-cellphone-locations-worldwide-snowden-documents-show/2013/12/04/5492873a-5cf2-11e3-bc56-c6ca94801fac_story.html.
[44] James Ball, *NSA Stores Metadata of Millions of Web Users for up to One Year*, GUARDIAN (UK), Sep. 30, 2013, http://www.theguardian.com/world/2013/sep/30/nsa-americans-metadata-year-documents.
[45] In Re Application of the Federal Bureau of Investigations for an Order Requiring the Production of Tangible Things from Verizon Business Networks Inc. on Behalf of MCI Cummunications Services, Inc. D/B/A/ Verizon Business Services, BR 13-80 (F.I.S.C. Apr. 15, 2013) [hereinafter "*Verizon Order*"], *available at* http://www.theguardian.com/world/interactive/2013/jun/06/verizon-telephone-data-court-order.
[46] Pema Levy, *NSA FISA Metadata Surveillance: Is the Government Using Cell Phones to Gather Location Data?*, INT'L BUSINESS TIMES, Aug. 2, 2013, http://www.ibtimes.com/nsa-fisa-metadata-surveillance-government-using-cell-phones-gather-location-data-1370221. Of course, any decision invalidating the NSA's surveillance programs would have likely been appealed to the Supreme Court.
[47] Klayman v. Obama, 13-cv-00851 (D.D.C. Dec. 16, 2013)
[48] *Verizon Order, supra* note 45.
[49] Barton Gellman, *NSA Broke Privacy Rules Thousands of Times Per Year, Audit* Finds, WA. POST, Aug. 15, 2013, http://www.washingtonpost.com/world/national-security/nsa-broke-privacy-rules-thousands-of-times-per-year-audit-finds/2013/08/15/3310e554-05ca-11e3-a07f-49ddc7417125_story.html.
[50] Brian Fung, *The NSA Is Giving Your Phone Records to the DEA and Covering It Up*, WA. POST, Aug. 5, 2013, http://www.washingtonpost.com/blogs/the-switch/wp/2013/08/05/the-nsa-is-giving-your-phone-records-to-the-dea-and-the-dea-is-covering-it-up/.
[51] Andrea Peterson, *LOVEINT: When NSA Officers Use Their Spying Power on Love Interests*, WA. POST, Aug. 24, 2013, http://www.washingtonpost.com/blogs/the-switch/wp/2013/08/24/loveint-when-nsa-officers-use-their-spying-power-on-love-interests/.
[52] Charlie Savage & Leslie Kaupman, *Phone Records of Journalists Seized by U.S.*, N.Y. TIMES, May 14, 2013, http://www.nytimes.com/2013/05/14/us/phone-records-of-journalists-of-the-associated-press-seized-by-us.html.
In the same investigation, the government did obtain the contents of emails of one reporter by obtaining a search warrant – but there is no indication that a warrant was used to seize the calling records themselves.
[53] Chris Soghoian, *8 Million Reasons for Real Surveillance Oversight*, SLIGHT PARANOIA, Dec. 1, 2009, http://paranoia.dubfire.net/2009/12/8-million-reasons-for-real-surveillance.html.
[54] *See Cell Phone Tracking Public Records Request*, https://www.aclu.org/protecting-civil-liberties-digital-age/cell-phone-location-tracking-public-records-request.
[55] *E.g.*, *Google Transparency Report*, http://www.google.com/transparencyreport/; *Twitter Transparency Report*, https://blog.twitter.com/2012/twitter-transparency-report.
[56] *Responses Received from Wireless Carriers on Law Enforcement Requests*, http://www.markey.senate.gov/Markey_Receives_Responses_from_Wireless_Carriers_on_Law_Enforcement_Requests.cfm.
[57] James Ball, *US Drug Agency Partners with AT&T for Access to "Vast Database" of Call Records*, GUARDIAN (UK), Sep. 2, 2013, http://www.theguardian.com/world/2013/sep/02/nsa-dea-at-t-call-records-access.
[58] YOU ARE BEING TRACKED: HOW LICENSE PLATE READERS ARE BEING USED TO RECORD AMERICANS' MOVEMENT 8, https://www.aclu.org/alpr.
[59] John Kelly, *Cellphone Spying: It's Not Just the NSA*, USA TODAY, http://www.usatoday.com/story/news/nation/2013/12/08/cellphone-data-spying-nsa-police/3902809/.
[60] *See, e.g.*, Sean Gallagher, *Staking Out Twitter and Facebook, New Service Lets Police Poke Perps*, ARSTECHNICA, Nov. 13, 2013, http://arstechnica.com/information-technology/2013/11/staking-out-twitter-and-facebook-new-service-lets-police-poke-perps/.
[61] Michael Isikoff, *The Snitch in Your Pocket*, NEWSWEEK, Feb. 19, 2010, http://www.newsweek.com/2010/02/18/the-snitch-in-your-pocket.html.
[62] *See id*; Comments of Al Gidari, Where I'm Calling From, On the Media (NPR radio broadcast May 8, 2009), *available at* http://www.onthemedia.org/transcripts/2009/05/08/05.
[63] Kelly, *supra* note 59.
[64] McIntyre v. Ohio Elections Commission, 514 U.S. 334 (1995).
[65] *Id.*; Watchtower Bible and Tract Society of New York, Inc., et al., v. Village of Stratton, et al., 536 U.S. 150 (2002).
[66] Daniel Solove, *First Amendment as Criminal Procedure*, 82 N.Y. UNIV. LAW REV. 112, 145 (2007).
[67] It is unclear whether this same degree of protection applies in the context of a criminal investigation, where at least one court has held that the identity of an anonymous speaker can be compelled unless there is "no reasonable possibility" that the information will be relevant to the prosecution of a crime. United States v. R. Enters., Inc., 498 U.S. 292, 301, 111 S. Ct. 722, 728, 112 L. Ed. 2d 795 (1991). However, in other instances law enforcement has backed down

Electronic copy available at: https://ssrn.com/abstract=2573962

from efforts to compel information about the identity of an online speaker. *See, e.g.,* Elinor Mills, *Twitter Challenges Court Order to Hand over User Data*, CNET, May 8, 2012, http://news.cnet.com/8301-1009_3-57430273-83/twitter-challenges-court-order-to-hand-over-user-data/.

[68] *See* Mills, *supra* note 67.

[69] United States v. Rumely, 345 U.S. 41 (1953).

[70] In Re Grand Jury Subpoena to Kramerbooks and Afterwords, Inc., 26 Media L. Rep. 1599 (D.D.C. 1998).

[71] Tattered Cover, Inc. v. City of Thornton, 44 P.3d 1044 (Colo. 2002).

[72] Amazon. com LLC v. Lay, 758 F. Supp. 2d 1154, 1167 (W.D. Wa. 2010).

[73] Video Privacy Protection Act, 18 U.S.C. § 2710 (2012).

[74] California Reader Privacy Act. CA. CIVIL CODE §1798.80 (2012).

[75] American Library Ass'n, *State Privacy Laws Regarding Library Records*, http://www.ala.org/offices/oif/ifgroups/stateifcchairs/stateifcinaction/stateprivacy.

[76] National Association for the Advancement of Colored People v. Alabama, 357 U.S. 449 (1958).

[77] *See ACLU v. Clapper — Challenge to NSA Mass Call-Tracking Program*, https://www.aclu.org/national-security/aclu-v-clapper-challenge-nsa-mass-phone-call-tracking; Electronic Frontier Foundation, *First Unitarian Church of Los Angeles v. NSA*, https://www.eff.org/cases/first-unitarian-church-los-angeles-v-nsa.

[78] United States v. Jones, 565 US ___, 132 S.Ct. 945, 962–64 (2012) (Alito, J. Ginsberg, J., Breyer, J. and Kagan, J., concurring in the judgment).

[79] *Id*. at 957 (Sotomayor, J., concurring).

[80] State v. Earls, A-53-11 (N.J. July 18, 2013).

[81] Commonwealth v. Rousseau, SJC-11227 (Ma. Supreme Judicial Ct. June 5, 2013).

[82] *See* Allie Bohm, *First in the Nation: Montana Requires a Warrant for Location Tracking*, FREE FUTURE, June 20, 2013, https://www.aclu.org/blog/technology-and-liberty-national-security/first-nation-montana-requires-warrant-location.

[83] *See* Alanna Durkin, *New Maine Cellphone Privacy Laws Take Effect Wednesday*, PORTLAND PRESS HERALD, Oct. 8, 2013, http://www.pressherald.com/news/New_cellphone_privacy_laws_take_effect_Wednesday_.html.

[84] Somini Sengupta, *With Montana's Lead, States May Demand Warrants for Cellphone Data*, N.Y. TIMES BITS, July 2, 2013, http://bits.blogs.nytimes.com/2013/07/02/with-montanas-lead-states-may-demand-warrants-for-cellphone-data/.

[85] *See Geolocation Privacy Legislation*, http://www.gps.gov/policy/legislation/gps-act/.

[86] Klayman v. Obama, 13-cv-00851 (D.D.C. Dec. 16, 2013), *available at* http://legaltimes.typepad.com/files/obamansa.pdf. However, in another December 2013 case, a N.Y. district court judge ruled that the program was constitutional. American Civil Liberties Union v. Clapper, 13 Civ. 3994 (WHP) (S.D. N.Y. Dec. 27, 2013) (Memorandum & Order).

[87] *See ACLU v. Clapper — Challenge to NSA Mass Call-Tracking Program*, https://www.aclu.org/national-security/aclu-v-clapper-challenge-nsa-mass-phone-call-tracking.

[88] *See* Gaurav Laroia, *The USA FREEDOM Act Answers Judge Leon's Constitutional Concerns*, BLOG OF RIGHTS, Dec. 18, 2013, https://www.aclu.org/blog/national-security/usa-freedom-act-answers-judge-leons-constitutional-concerns.

[89] Transcript of President Obama's Jan. 17 Speech on NSA Reform, WA. POST, Jan. 17, 2014, http://www.washingtonpost.com/politics/full-text-of-president-obamas-jan-17-speech-on-nsa-reforms/2014/01/17/fa33590a-7f8c-11e3-9556-4a4bf7bcbd84_story.html.

[90] *See* Zeke J. Miller, *White House Group Calls for Limits on NSA Surveillance*, TIME SWAMPLAND, Dec. 18, 2013, http://swampland.time.com/2013/12/18/white-house-group-calls-for-limits-on-nsa-surveillance/.

[91] LIBERTY AND SECURITY IN A CHANGING WORLD, REPORT AND RECOMMENDATIONS OF THE PRESIDENT'S REVIEW GROUP ON INTELLIGENCE AND COMMUNICATIONS TECHNOLOGIES, Dec. 12, 2013, http://www.whitehouse.gov/sites/default/files/docs/2013-12-12_rg_final_report.pdf.

[92] PRIVACY & CIVIL LIBERTIES OVERSIGHT BOARD, REPORT ON THE TELEPHONE RECORDS PROGRAM CONDUCTED UNDER SECTION 215 OF THE USA PATRIOT ACT AND ON THE OPERATIONS OF THE FOREIGN INTELLIGENCE SURVEILLANCE COURT, Jan. 23, 2014, http://www.pclob.gov/SiteAssets/Pages/default/PCLOB-Report-on-the-Telephone-Records-Program.pdf.

[93] United States v. Jones, 565 US _, 132 S.Ct. 945, 957 (2012) (Sotomayor, J., concurring).

[94] United States v. Powell, _ F.Supp.2d _, 2013 WL 1876761 (E.D. Mich May 3, 2013).

[95] Klayman v. Obama, 13-cv-00851 (D.D.C. Dec. 16, 2013).

[96] *See* Burrows v. Superior Court, 529 P.2d 590, 593 (Cal. 1974). In principle, these decisions should preclude warrantless demands for metadata by state and local government actors. As a practical matter, however, their impact has remained limited; for example, law enforcement officials in states such as California continue to acquire location information and other metadata without a search warrant.

[97] United States v. Maynard, 615 F.3d 544, 561–63 (D.C. Cir. 2010), aff'd sub nom. United States v. Jones, 565 US ___, 132 S.Ct. 945 (2012).

[98] *See, e.g.*, Orin Kerr, *The Mosaic Theory of the Fourth Amendment*, 111 Mich. L. Rev. 3 (2012).

[99] Optiver Australia Pty. Ltd. & Anor. V. Tibra Trading Pty. Ltd. & Ors., No. C 12-80242 EJD (PSG), 2013 WL 256771 (N.D. Cal. Jan. 23, 2013) at *1.

[100] LIBERTY AND SECURITY IN A CHANGING WORLD, REPORT AND RECOMMENDATIONS OF THE PRESIDENT'S REVIEW GROUP ON INTELLIGENCE AND COMMUNICATIONS TECHNOLOGIES, Dec. 12, 2013, http://www.whitehouse.gov/sites/default/files/docs/2013-12-12_rg_final_report.pdf.

[101] *See Wiretap Report FAQs*, http://www.uscourts.gov/Statistics/WiretapReports/wiretap-report-faqs.aspx.

Electronic copy available at: https://ssrn.com/abstract=2573962

"THE GOVERNMENT SURVEILLANCE PROGRAMS CONDUCTED UNDER THE FOREIGN SURVEILLANCE INTELLIGENCE ACT ARE FAR BROADER THAN THE AMERICAN PEOPLE PREVIOUSLY UNDERSTOOD. IT IS TIME FOR SERIOUS AND MEANINGFUL REFORMS SO WE CAN RESTORE CONFIDENCE IN OUR INTELLIGENCE COMMUNITY."

–REP. JIM SENSENBRENNER (R-WI)

"WE RECOMMEND THAT THE GOVERNMENT SHOULD COMMISSION A STUDY OF THE LEGAL AND POLICY OPTIONS FOR ASSESSING THE DISTINCTION BETWEEN META-DATA AND OTHER TYPES OF INFORMATION."

–PRESIDENT'S REVIEW GROUP ON INTELLIGENCE AND COMMUNICATIONS TECHNOLOGIES

"THERE'S NOT A SHARP DIFFERENCE BETWEEN METADATA AND CONTENT.... IT'S MORE OF A CONTINUUM."

–MICHAEL MORELL, FORMER CIA OFFICIAL

Online at www.aclunc.org/tech/meta